Proceedings of ANAS (Biological and Medical Sciences), vol. 73, No 2, p. 156-163 (2018)

Creation and Development of the Information System of the Plant Genepool of Azerbaijan

Z.I. Akparov, I.A. Mirzaliyeva, A.T. Mammadov*

*Genetic Resources Institute, Azerbaijan National Academy of Sciences, 155 Azadlig Ave., Baku AZ1106, Azerbaijan; *E-mail: afig.mammadov@gmail.com*

The Information System that provides optimal monitoring of the plant genetic resources and enables us utilize the research findings regarding the spread, preservation and morpho-agronomic, biogeographic, technological, biochemical and molecular traits, genetic diversity and modifications of crops and crop wild relatives in Azerbaijan in the process of collection, storage, analysis, breeding and in the other scientific works in a centralized space and standardized form, as well as a number of databases covering various traits of domestic plant gene pools, and the bases of characterization and evaluation of 13 species of grain cereals and leguminuous crops conserved in the National GeneBank were created, developed and integrated into the international systems.

Keywords: *Information system, database, genetic resources, genepool, characterization, evaluation, plant*

INTRODUCTION

In the 80-90s of the last century, the notable Azerbaijani scientist of biology and agricultural sciences, academician Jalal Aliyev called attention to the necessity of the information systems for reliable conservation and sustainable use of biological diversity, raised concerns over creation of the information bases that cover passport and comprehensive research data of the plant genepool samplings and put forward his perspicacious recommendations. He stressed the need for creation of the information systems aimed at on time revealing of the overwhelming risks related to PGR and their root causes, simulating the expected development of resources, suggesting recommendations on the basis of the predictions, gathering information in a centralized and standardized manner for information provision of decision-making mechanisms. Later the thesis emphasizes that the information system creates opportunities for studying and efficient management of the bioresources were put forth number of times (Thysen, 2000; Guarino et al., 2002; Əliyev və b., 2008; Descriptors – BI). The related priorities specified by J.Aliyev were systemized afterwards into an appropriate action plan of which key priorities were reflected in the national strategies and state programs (Əliyev, Əkpərov, 2002).

In the study of genetic resources that are the carriers of informative aspects of animate nature, are applied the methods typical for the field, as well as informational-analytical software programs and databases based on standardization of quality and quantity indicators (Наумов, Вендров, 1991; Thysen, 2000; Winfried, 2006).

Special descriptors (Germier, Frese, 2001; Descriptors – BI, 1982-2018; FAO/BI Multi-Crop Passport Descriptors list, 2015) developed by FAO and Global Biodiversity are applied to study genetic diversity and modification and morpho-agronomic, biogeographical and so forth traits (Guarino et al., 2002; Akparov et al., 2013), whereas phenotypical, metabolic, biochemical and molecular characterizations are performed via various markers (RFLP, protein fractions, microsatellites, RAPD, SSR, EST, SNP). It is tremendously necessary to organize such researches as a database so that they can be digitalized and collected in a centralized space and standard forms, stored, manipulated and more efficiently used in the other researches (Thysen, 2000; Germier, Frese, 2001; Guarino et al., 2002; Winfried, 2006; Əliyev və b., 2008). Such databases of the germplasms enable us to automate analyses and selections, eliminate the need for multiple researches, and save funds and time.

Characterization of plant genetic resources is the major description of germplasm for breeding. Their morpho-agronomic, biochemistry and molecular evaluation also plays an important role for creation new valuable varieties. With the support of information technologies, it can be more effective use of characterization and evaluation data.

The researches we have conducted for long years have basically aimed at development of the bases of an information system that posseses all the foregoing traits, and creation of passport, taxonomic and ecological databases, as well as agronomic, characterization, evaluation, genomics ones on the basis of an enhanced software.

MATERIALS AND METHODS

The research materials of the national information system on PGR include passport, ecological, climatic, geological and geobotanical, areological and taxonomic storage, recovery, exchange, introduction and reintroduction, characterization and evaluation data standardized by relevant technologies with regard to the collections of the national genepool, scientific and national breeding varieties, wild crop relatives, herbarium funds and plant samplings. The objects of the system are the expeditions, reports and description lists on study of genepool materials, electronic maps, historical and archeological materials, initial registration and quarantine protocols, collectors, farmers, experts and their PGR-related knowledge, organizations, projects, programs, publications, catalogues, other references and materials (Акпаров, Мамедов и др., 2007; Akparov et al., 2008).

In the creation of the multi-table and multi-level databases were used the Database Management Systems and other program packages (Visual FoxPro, dBase, MS Access, MySQL, SQL Server, Apachi, Oracle, MS Excel və s.), other open-source software and database servers and internet resources, and SQL language (Structured Query Language) applied in built-in software writing (Наумов, Вендров, 1991; Сосински, 1997; Каратыгин и др., 1999; Дейт, 2005; Кузнецов, 2007; Germier, Frese, 2001; GENESYS; ECPGR Germplasm Databases).

The gathering, determination and analysis of the taxonomic data included into the system were put into practice through online version of the Index Kewensis system, GRIN-Taxonomy site, Vascular Plant Families and Genera, Mansfeld's World Database, other relevant internet resources and related databases (The Index Kewensis; GRİN Taxonomy; Mansfeld's World Database, 2012; Vascular Plant Families and Genera).

International passport and characterization descriptors (as well as genetic marker) were used for standardization, digitalization, collection and processing of research results for studying of plants (Descriptors – BI, 1982-2018; FAO/BI Multi-Crop Passport Descriptors list, 2015; GENESYS; ECPGR Germplasm Databases). In the analysis of the informational analysis of the national genepool were applied the quality factors and indexes on evaluation of genetic diversity of collections and ecosystems (Əliyev və b., 2008), whereas GIS software tools (Guarino et al., 2002; Winfried, 2006; Madurika, Hemakumara, 2017) and statistic processing packages of data were used in drawing up electronic areal maps of plant diversity, transferring and analysis of the ecogeographic, geobotanical and agronomic data and creation of relevant databases. The World Information and Early Warning System (WIEWS) of FAO, European Internet Search Catalogue – EURISCO on PGR and European Cooperative Program for Plant Genetic Resources (ECPGR Germplasm Databases), World Information Center for PGR (GENESYS) were used for integration into the international systems of the information system and work on the online mode.

Because the National Information Sharing Mechanism has the Web-based program, it needs a Web browser. High versions of Explorer or Mozilla Firefox etc. can be used. The software was created by using non-licensed, open source software (for example, Java Servlet and Java Server Pages), Web-server - Tomcat 4.1.24 and database server McKoi v.94.

RESULTS AND DISCUSSION

Functional blocks of the information system and creation of its fundamental structure. The biodiversity can be much better conserved by organization of electronic information resources on the biological systems in a different way and by securing information flow, in consideration of which were specified the key principles of the information flows that provide an efficient and comfortable access to resources, freedom of action in the rich information space for users, operativeness in information search and selection and inquiry, settlement of maintenance of actual condition of all information resources and provision of uninterrupted work of the documentation and information flow of genepools, and organizational structure and software were created. Such structure enables us to gather and analyze complex passport, morphological, botanical and ecnological data, including but not limited to characterization and evaluation and genomix data. The key principles of user-system and system-user interactice model were developed within the information system on PGR.

The Information System for PGR of Azerbaijan consists of the National Information Sharing Mechanism on PGRFA and its internet-based databases and other multiple file groups, multi-table and multifunctional Central Database of the national genepools, characterization and evaluation databases integrated into its structure and other software tools and relevant interfaces.

The Information System can be divided into three main functional blocks: 1) Inventarization (information gathering) block; 2) Analytical block; 3) Organizational-management block. Applying information technologies, all blocks have been closely associated with the software tools.

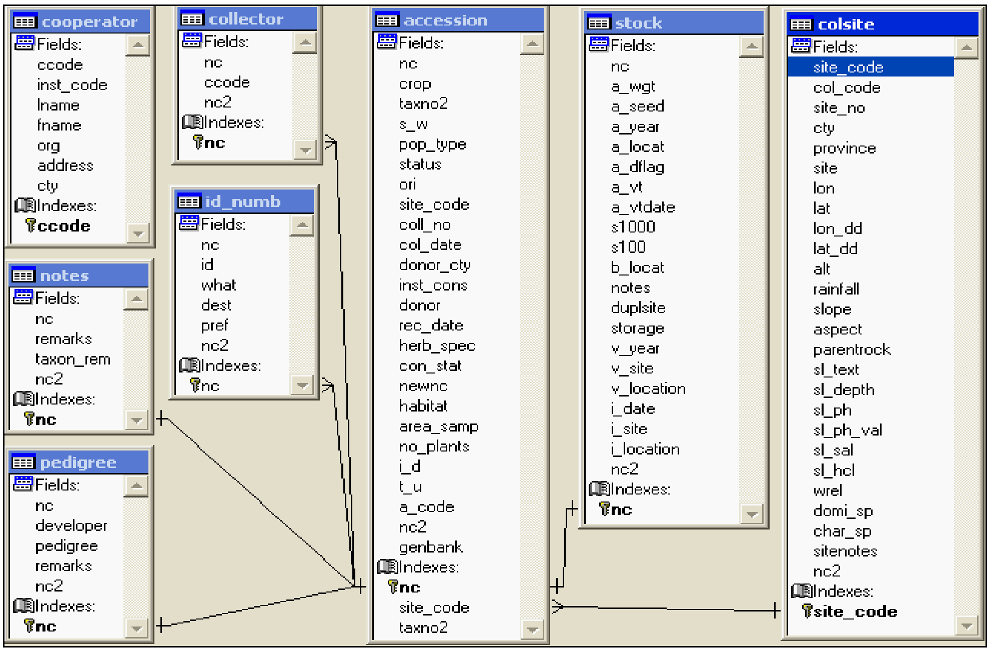


Fig. 1. A view of the internal structure of the database.

The Central Information Base mainly comprises of software made by modern technologies, and tables coordinated by special key fields and cover the most various issues of the PGR. The data on passport, ecological, botanical details and storage of the plant samplings have been systemized separately. Besides, the tables were projected and coordinated by coding and indexes to store additional information on subjects, taxon names and countries that collect and conserve plant samplings (Figure 1).

The Central Database can also be used for statistic analysis. The selections and surveys once made are saved and thus, no need arises to make such surveys once again.

Database of the National Genebank was established within the structure of the Central Database. The structure of the database that contains 14 tables, explanations and additional blocks, search and report forms adapted to the nature of the GeneBank activities has an efficient structure and interface to reach data consecutively and in a quick way, and provides better management of data describing the germplasm.

The internet-based and multi-table National Information Sharing Mechanism that secures high quality management of PGR research activities cover the data on *in situ* and *ex situ* conservation of PGR, their study, breeding and seed-growing. Through their research, regular strategic analyses are conducted on the condition of the PGR in Azerbaijan. Using keys in an online mode, it is possible to make searches and selections for all databases or their separate parts on the search system page.

The intellectual solutions blocks of the system we have created enable us to achieve a development plan of the condition on the basis of integral evaluation and expert system (knowledge database) of the prediction characteristics. The creation of the knowledge base built in accordance with the principles that enable us to protect and efficiently use the PGR create suitable condition for revealing, display and solution of the most serious problems of protection of plant biodiversity, and provide sustainable development of the genetic resources. The base acts on the basis of decision-making computerization, at the time of which creation there were taken into account expert answers to the questions on PGR in various forms.

Ecogeographical analyses within the information system. The organization of the data in geographic systems was believed to be necessary in studying of the space-time aspects of the development of plant diversity. Having applied the Central Database tools, the places where national genepool samplings are found, passport, characterization and evaluation data have been included into the distribution maps of the species. There were analyzed the possibility on availability of the materials collected upon addition of the climatic characteristics of the territory or the variation of the quality indicators in respect of the height above sea level. The genetic erosion of the country was evaluated on the basis of various data gathered to provide information support of the strategic planning and by specifying precisely nature and range of problems and the root of dangers for biodiversity.

In the 10 regions of the republic, a big amount of data that cover more than 200 farms and 678 points on local, traditional and national selection of

pomegranate, as a sample, were gathered and analyzed. It turned out to be possible to specify the diversity of national selection varieties, modification dynamics of diversity, condition of on-farm protection on lands and the distribution properties of the varieties. Besides, there were analyzed the relations among species diversity, absolute height, disease catching, local bioclimatic indicators and such other elements. The richness indexes of the Margalef (10,28), Shannon-Viner (2,972) and Simpson (0,901) proved the availability of variable diversity in different regions (Figure 2.).

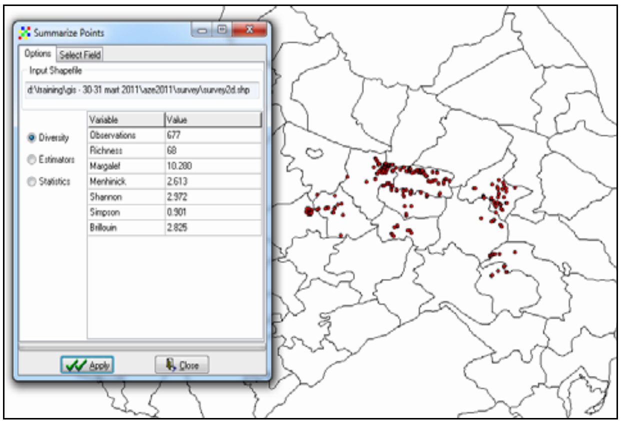


Fig. 2. Evaluation of the variable diversity in the regions of study.

Climatic data and other indicators prove that, the domestic pomegranate varieties have reached a respective adaptivity as a result of utilization of wild crops and multi-century breeding. According to the researches, in the territories where the amount of annual rainfall is higher, the rate of disease catching is high too.

The data received from the researches have been included into the relevant database together with the other agronomic data.

Content (accession, taxonomical) of the Central DB. Original crops, CWR, traditional and modern varieties are widely presented in the Central Database of which principal part constitute the data on cereals, leguminous and technical plants, fruits and berries of high priority.

There are more than 13500 plant accessions included in the database and conserved in the main *ex situ* collections of Azerbaijan until the present day. These accessions refer to 113 families, 443 genera, 871 species, and 304 variations. Of these, 5135 accessions are cereals, 1513 legumes, 751 fodders (forages), 873 vegetable-melons, 1586 technical plants, 2414 fruits, 786 medicines and etc. More than 7900 seed samples are in the medium-term storage cell.

According to the biological status of the registered samples from the CDB, 4861 were classified as modern scientific varieties, 2365 of the folk selection, 3652 constant research materials, 2205 CWR and etc.

Families Poaceae, Malvaceae, Fabaceae, Rosaceae, Vitaceae, Solanaceae, Lythraceae, Apiaceae, Cucurbitaceae, Betulaceae, Moraceae have been represented with more accessions in the central database of the *ex situ* collections of the Republic. The species *Gossypium hirsutum, Triticum aestivum, Zea mays, Vitis vinifera, Pyrus communis, Triticum status, Malus domestica, Hordeum vulgare, Phaseolus vulgaris, Punica granatum, Corylus avellana, Gossypium barbadense, Medicago sativa, Ficus carica,* Cicer *arietinum* differ with more accessions.

Characterization – evaluation databases. The characterization bases of the wheat, barley, triticale, rye, bean, cowpea, vetch, pea, lentil, broad bean and grass pea (13 species, 2186 accessions) were created in the structure of the Central Database via FoxPro program (Table 1).

A fragment of the small programs used in creation of such databases is given below:

Table. Species and accession content of characterization and evaluation databases

N	Botanical name of the species	Common name of the plant	Number of evaluated accessions	Number of traits for evaluation
1.	*Lathyrus sativus* L.	Grass pea	67	25
2.	*Phaseolus vulgaris* L.	Green bean	93	18
3.	*Lens culinaris* Medic.	Lentil	85	34
4.	*Cicer arietinum* L.	Chickpea	209	34
5.	*Zea mays* L.	Maize	177	10
6.	*Vicia faba* L.	Broad bean	89	24
7.	*Vigna unungiucullata* (L.)Walp.	Cowpea	25	18
8. 9.	*Triticum aestivum* L. *Triticum durum* Desf.	Common and durum wheat	1033	10
10.	*Hordeum vulgare* L.	Barley	194	22
11.	*Triticale*	Triticale	19	12
12.	*Secale*	Rye	135	12
13.	*Vicia*	Vetch	60	24
		Total	**2186**	

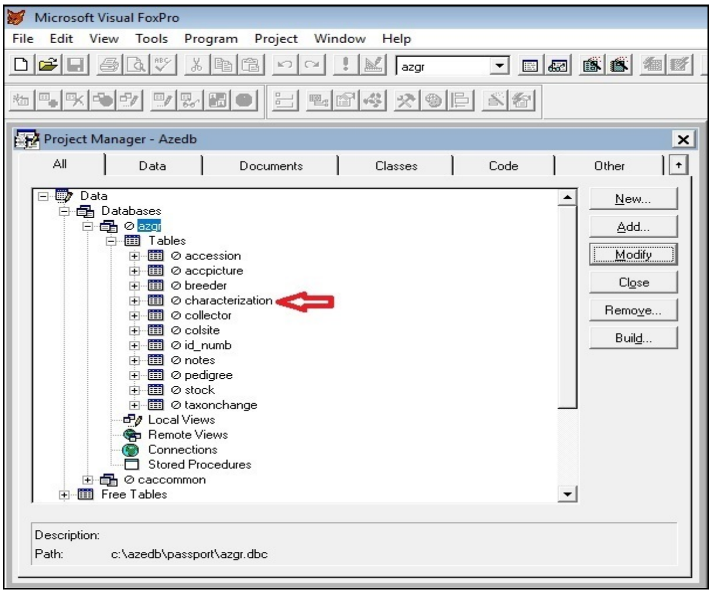


Fig. 3. Description of the characterization table in the given database.

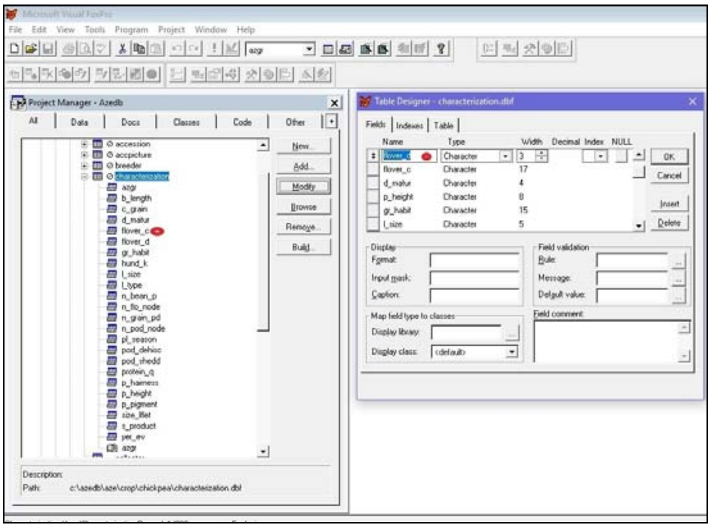


Fig. 4. Fields of table that makes the characterization database.

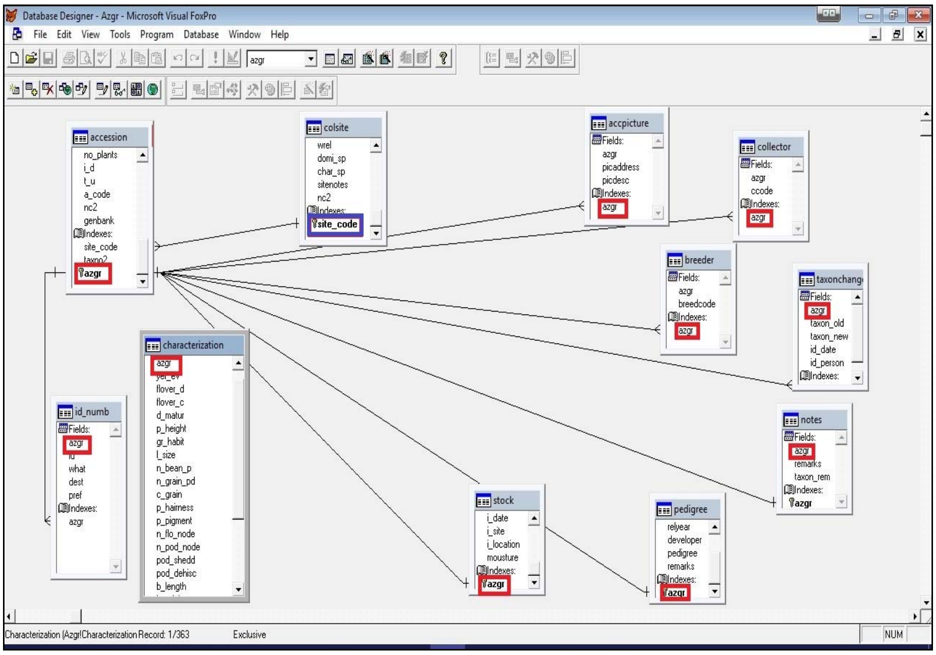


Fig. 5. Coordination of the other tables per AzGR attributive, including characterization table of the National Database.

AZE: wheat, wheat accession in Azerbaijan collection, C:\cac\AZE\cropsel\wheat\SELECTIONS; C:\cac\AZE\crop\wheat\DB_CROP.dbf C:\cac\AZE\cacsel\db_user.dbf; SELECT nc FROM accession WHERE taxno2 IN (SELECT taxno2 FROM taxon2 WHERE atc ('wheat',tax_name) > 0.

"CHARACTERIZATION" tables of different structure per plants were created in every database to store characterization and evaluation data (Fig. 3, 4). The relations between all tables that make up the base have been built by national AzGR indicator (Fig. 5).

Under guidance of the breeders and experts, the appropriate characterization and evaluation databases of separate plants and plant groups take a defined shape and are developed on the basis of morphological, physiological, technological, biochemical immunological and genetic study of cereals and leguminous plant collections of the National GeneBank.

CONCLUSIONS

There were created, for the first time, the Information System on PGR of Azerbaijan that provides the gathering, systemization, storage, spread, exchange and optimal management of the data on plant genetic resources, the Central Database that is an integral part of the Information System and the other databases that cover the data space on such resources in different directions. There were established the National Information Exchange Mechanism that provides coordination of movement dynamics and management of PGR, and the database that helps decision-making mechanisms on regular monitoring of the condition of genepool and its sustainable use.

In Azerbaijan, the relations among the spread and agronomic indicators of the national breeding varieties, adaptability of varieties and forms and agroclimatic gradient were studied for the first time in the background of the bioclimatic data of the regions where pomegranate, as a sample, is grown, and the diversity of which conservation is the highest priority was ascertained.

There were created characterization and evaluation databases of wheat, barley, maize, triticale, rye, cowpea, chickpea, lentil, broad bean and grass pea conserved in the National GeneBank.

Created database system is an open access for country and Institutional users (breeders, scientists, PhD students, specialists), and in near future will be online like a web-based Information system for all users in the world.

REFERENCES

Əliyev C.A., Əkpərov Z.İ. (2002) Azərbaycanın bitki genetik ehtiyatları. AMEA-nın Xəbərləri (biologiya elmləri seriyası), **1-6:** 57-68

Əliyev C.Ə., Əkpərov Z.İ., Məmmədov A.T. (2008) Bioloji Müxtəliflik. Bakı: Elm, 232 s.

Ərzaq və Kənd Təsərrüfatı Bitki Genetik Ehtiyatları üzrə Milli İnformasiya Mübadiləsi Mexanizmi. MİMM (2006-2018) (İnternet portalı: http://www.fao.org/pgrfa-gpa-archive/aze/welcome.htmx)

Акпаров З.И., Мамедов А.Т. (2007) Информационная система по генетическим ресурсам растений Азербайджана. *Современные проблемы науки и образования* (Москва), **№6**: 9-13.

Кузнецов С.Д. (2007) Основы баз данных, 2-е изд., М.: Интернет-университет информационных технологий; БИНОМ, Лаборатория знаний, 484 с.

Дейт К.Дж. (2005) Введение в системы баз данных [Introduction to Database Systems]. 8-е изд. М.: Вильямс, 1328 с.

Каратыгин С., Тихинов А., Тихинова Л. (1999) Visual FoxPro 6. ЗАО «Издательство БИНОМ».

Наумов А.Н., Вендров А.М. и др. (1991) Системы управления базами данных и знаний. Москва: Финансы и статистика, 343 с.

Сосински Б. (1997) Разработка приложений в среде Visual FoxPro 5. Пер. с анг. Киев: Диалектика, 448 с.

Akparov Z., Mammadov A., Mirzaliyeva I. (2008) Database of the Azerbaijan National Genebank. *Bulletin of the Academy of Agricultural Sciences of Georgia* (Tbilisi), p. 165-166.

Akparov Z.I., Mammadova S.A., Mammadov A.T. (2013) Mobilization and conservation of the seed pool of plant genetic resources in Azerbaijan. *Russian Journal of Genetics: Applied Research*, **3(1):** 26-29.

Descriptors – Bioversity International (https://www.bioversityinternational.org/e-library/publications/descriptors/)

ECPGR Germplasm Databases (http://www.ecpgr.cgiar.org/resources/germplasm-databases/)

FAO/BI Multi-Crop Passport Descriptors list, (2015) https://www.bioversityinternational.org/e-library/publications/detail/faobioversity-multi-crop-passport-descriptors-v21-mcpd-v21/

GENESYS – the Global Gateway to Genetic Resources (https://www.genesys-pgr.org)

Germier Ch.U., Frese L. (2001) A data model for the evaluation and characterization of plant genetic resources. *Broad Variation and Precise Characterization – Limitation for the Future*. Poland: Poznan, p. 174-177.

GRIN Taxonomy (https://npgsweb.ars-grin. gov/ gringlobal/taxon/taxonomysearch.aspx)

Guarino L, Jarvis A, Hijmans RJ, Maxted N. (2002) Geographic information systems (GIS) and the conservation and use of plant genetic resources. In: Engels J.M.M., Ramantha R.V., Brown A.H.D., Jackson M.T. (eds.): *Managing Plant Genetic Diversity*. UK, Wallingford" CABI Publishing, p. 387-404.

Madurika H.K.G.M., Hemakumara G.P.T.S. (2017) GIS Based Analysis For Suitability Location Finding In The Residential Development Areas Of Greater Matara Region. *International Journal of Scientific & Technology Research*, **4(8):** 96-105.

Mansfeld's World Database of Agricultural and Horticultural Crops (http://mansfeld.ipk-gatersleben.de/apex/f?p=185:3:0:)

The New Plant Genetic Resource Gateway: GENESYS (that currently compiles the data from SINGER, EURISCO and GRIN) (http://www. genesys-pgr.org/)

The Index Kewensis - The International Plant Names Index (https://www.ipni.org)

Thysen (2000) Agriculture in the information society. *J. Agric. Eng. Res.*, **76:** 297-303.

WIEWS - World Information and Early Warning System on Plant Genetic Resources for Food and Agriculture (http://www.fao.org/wiews/en/)

Winfried S. (2006) GIS, geostatistics, metadata banking, and tree-based models for data analysis and mapping in environmental monitoring and epidemiology. *International Journal of Medical Microbiology*, **296(S1):** p. 23-36.

Vascular Plant Families and Genera (Ali Bitkilərin Fəsilə və Cins Adları Siyahısı), (www.kew.org/web.dbs/genlist.html)

Azərbaycanın Bitki Genefondu İnformasiya Sisteminin Yaradılması Və İnkişafı

Z.İ. Əkpərov, İ.A. Mirzəliyeva, A.T. Məmmədov

AMEA Genetik Ehtiyatlar İnstitutu

Azərbaycanın mədəni bitkilərinin və onların yabanı əcdadlarının yayılması, mühafizəsi, morfo-aqronomik, biocoğrafi, texnoloji, biokimyəvi və molekulyar xüsusiyyətləri, genetik müxtəlifliyi və dəyişkənliyi üzrə tədqiqat nəticələrinin vahid məkanda və standart formalarda toplanması, saxlanması, təhlili, seleksiya işlərində və digər elmi işlərdə daha səmərəli istifadəsinə imkan verən, bitki genetik ehtiyatlarının optimal idarə olunmasını təmin edən İnformasiya Sistemi və onun strukturunda milli bitki genefondunun müxtəlif cəhətlərini əhatə edən çoxsaylı məlumat bazaları (o cümlədən, Milli Genbankda mühafizə edilən 13 dənli taxıl və paxlalı bitki növü üzrə səciyyələndirmə və qiymətləndirmə bazaları) yaradılmış və beynəlxalq sistemlərə inteqrasiya edilmişdir.

Açar sözlər: *İnformasiya sistemi, məlumat bazası, genetik ehtiyatlar, genefond, səciyyələndirmə, qiymətləndirmə, bitki*

Создание И Развитие Информационной Системы Генофонда Растений Азербайджана

З.И. Акпаров, И.А. Мирзалиева, А.Т. Мамедов

Институт генетических ресурсов НАН Азербайджана

Создана и интегрирована в международные системы Информационная Система и входящие в её структуру многочисленные базы данных, охватывающие различные аспекты национального генофонда (в том числе, характеристические и оценочные базы 13 видов зерновых и зернобобовых растений, сохраняемых в Национальном Генбанке), обеспечивающие оптимальное управление генетическими ресурсами растений, позволяющие в едином информационном пространстве и в стандартных формах сбор, хранение, анализ, рациональное использование в селекционных работах и в других научных работах данных по результатам исследований по распространению, защите, морфоагрономическим, биогеографическим, технологическим, биохимическим и молекулярным характеристикам, по генетическому разнообразию и изменчивости культурных растений Азербайджана и их диких сородичей

Ключевые слова: *Информационная система, база данных, генетические ресурсы, генофонд, характеристика, оценка, растение*